

Analisis Menggunakan *Association Rule* dengan Algoritma Apriori pada Data Transaksi Retail

Andri Khoirul Manaf^{1*}, Febrian Maulana², Nabilah Junita Eka Putri³, Whereyga Putra Pradana⁴

¹Teknik Informatika/Fakultas
Ilmu Komputer
(Universitas Duta Bangsa
Surakarta)

^{1*}230103050@mhs.udb.ac.id

²Teknik Informatika/Fakultas
Ilmu Komputer
(Universitas Duta Bangsa
Surakarta)

²230103060@mhs.udb.ac.id

³Teknik Informatika/Fakultas
Ilmu Komputer
(Universitas Duta Bangsa
Surakarta)

³230103072@mhs.udb.ac.id

⁴Teknik Informatika/Fakultas
Ilmu Komputer
(Universitas Duta Bangsa
Surakarta)

⁴230103080@mhs.udb.ac.id

Abstrak— Perkembangan bisnis retail menghasilkan data transaksi dalam jumlah besar yang dapat dimanfaatkan untuk memperoleh informasi mengenai pola pembelian pelanggan. Namun, data tersebut sering kali belum dimanfaatkan secara optimal sebagai dasar pengambilan keputusan. Penelitian ini bertujuan menerapkan metode *Association Rule* menggunakan algoritma Apriori untuk menemukan hubungan antarproduk yang sering dibeli secara bersamaan serta mengembangkan aplikasi berbasis Streamlit sebagai media analisis. Dataset yang digunakan merupakan data transaksi retail yang terdiri atas 38.765 baris data dan setelah proses preprocessing diperoleh 3.898 transaksi valid dengan 167 jenis produk. Proses analisis menggunakan parameter minimum support sebesar 8%, minimum confidence sebesar 50%, minimum lift sebesar 1, dan maximum rule length sebanyak 3 item. Hasil penelitian menunjukkan bahwa sistem berhasil menghasilkan 65 frequent itemset dan 11 association rule yang memenuhi seluruh parameter. Aturan terbaik diperoleh pada kombinasi other vegetables dan rolls/buns menuju whole milk dengan nilai support sebesar 8,21%, confidence sebesar 55,94%, dan lift sebesar 1,221. Nilai tersebut menunjukkan adanya hubungan asosiasi positif sehingga informasi yang diperoleh dapat dimanfaatkan untuk strategi product bundling, *cross-selling*, penataan tata letak produk, serta pengelolaan persediaan barang. Aplikasi yang dikembangkan juga mampu menampilkan hasil analisis, visualisasi network graph, rekomendasi produk, dan fasilitas ekspor laporan sehingga dapat membantu pengambilan keputusan bisnis berbasis data.

Kata kunci— Association Rule, Algoritma Apriori, Data Mining, Market Basket Analysis, Retail.

Abstract— The rapid growth of the retail industry has generated a large volume of transaction data that can be utilized to discover customer purchasing patterns and support business decision-making. However, such data are often underutilized due to the lack of effective analytical methods. This study aims to implement the Association Rule technique using the Apriori algorithm to identify relationships among products frequently purchased together and to develop a Streamlit-based web application for interactive data analysis. The dataset consists of 38,765 transaction records, which were preprocessed into 3,898 valid transactions containing 167 unique products. The Apriori algorithm was applied using a minimum support of 8%, a minimum confidence of 50%, a minimum lift of 1.0, and a maximum rule length of three items. The experimental result produced 65 frequent itemsets and 11 association rules that satisfied all predefined thresholds. The strongest association rule was identified between other vegetables and rolls/buns leading to whole milk, with a support value of 8.21%, a confidence value of 55.94%, and a lift value of 1.221, indicating a positive association between the products. Furthermore, the developed application successfully provides interactive visualization, association rule analysis, product recommendation features, and report export functionality. The findings can support retail businesses in implementing product bundling, cross-selling strategies, shelf layout optimization, and inventory management to improve operational efficiency and increase sales performance.

Keywords— Association Rule, Apriori Algorithm, Data Mining, Market Basket Analysis, Retail Analytics.

I. PENDAHULUAN

Perkembangan teknologi informasi telah membawa dampak yang sangat signifikan pada berbagai sektor bisnis, khususnya dalam industri

perdagangan ritel [1], [2], [3]. Aktivitas transaksi penjualan harian di sektor ritel baik pada skala minimarket, kedai kopi, maupun e-commerce menghasilkan volume data yang sangat besar dan terus bertambah [4], [5], [6].

Meskipun demikian, tumpukan data mentah tersebut seringkali hanya menjadi arsip dan tidak memberikan nilai informasi yang berguna apabila tidak diolah menggunakan metode komputasi yang tepat [7], [8], [9]. Oleh karena itu, pelaku bisnis dituntut untuk dapat menganalisis data transaksi penjualan secara mendalam guna memahami perilaku, preferensi, serta kebiasaan belanja konsumen [2], [10]. Pemahaman mengenai pola pembelian konsumen ini menjadi kunci strategis bagi perusahaan untuk menyusun kampanye pemasaran, mengoptimalkan tata letak barang, mengelola stok inventori, serta merancang strategi cross-selling atau rekomendasi produk yang akurat [1], [11], [12].

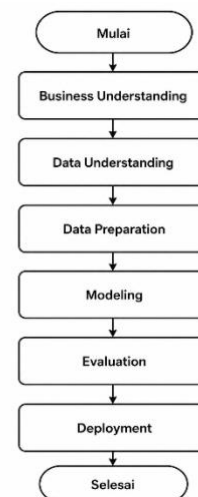
Untuk mengekstraksi informasi berharga dari basis data transaksi, pendekatan data mining yang paling efektif dan umum digunakan adalah Market Basket Analysis (MBA) [4], [6], [13]. Secara spesifik, teknik association rule mining di dalam MBA bertujuan untuk menemukan probabilitas dan pola hubungan antar beberapa item produk yang sering dibeli secara bersamaan oleh pelanggan di dalam satu keranjang belanja [8], [9], [14]. Dari berbagai algoritma penggalian aturan asosiasi, algoritma Apriori dinilai sebagai metode yang paling populer, andal, dan efisien dalam membentuk aturan asosiasi pada basis data berskala besar [9], [11]. Algoritma Apriori terbukti mampu memberikan tingkat akurasi yang tinggi dalam menggali pola kombinasi pembelian produk bila dibandingkan dengan algoritma lainnya [4], [10], [15]. Dalam implementasinya, algoritma ini beroperasi dengan mengevaluasi nilai minimum dukungan (support) untuk mengukur proporsi kemunculan kombinasi item, serta menggunakan nilai kepastian (confidence) untuk menguji kekuatan hubungan antar item tersebut [3], [6], [7].

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk mengimplementasikan teknik association rule menggunakan algoritma Apriori secara mendalam pada data transaksi ritel. Hasil dari analisis ini difokuskan untuk menemukan

himpunan produk yang paling sering dibeli secara bersamaan, sehingga menghasilkan aturan asosiasi yang kuat [5], [8]. Temuan penelitian ini diharapkan dapat memberikan rekomendasi yang berbasis data konkret (*data-driven*) guna membantu pihak manajemen ritel dalam pengambilan keputusan bisnis, seperti merancang paket promosi (*bundling*), mengoptimalkan penempatan rak produk, serta menunjang penawaran produk tambahan (*crossselling*) demi mendorong profitabilitas penjualan [11], [14].

II. METODOLOGI PENELITIAN

Penelitian dilakukan melalui tahapan-tahapan terstruktur berdasarkan kerangka kerja *Cross Industry Standard Process for Data Mining* (CRISP-DM), mulai dari pemahaman data, persiapan data, pemodelan hingga evaluasi hasil analisis.



Gambar 1. Tahapan Metode CRISP-DM

A. Business Understanding

Pada tahap *Business Understanding*, penelitian bertujuan untuk memanfaatkan data transaksi penjualan ritel guna menemukan pola pembelian konsumen menggunakan algoritma Apriori. Permasalahan yang dihadapi adalah besarnya data transaksi yang belum dimanfaatkan secara optimal sehingga belum mampu memberikan informasi mengenai keterkaitan antarproduk.

Hasil analisis diharapkan mampu membantu pihak retail dalam:

- menentukan strategi *cross-selling*,
- menyusun paket promosi (*bundling*),
- mengoptimalkan tata letak produk,
- membantu pengendalian stok barang.

Tahapan ini menjadi dasar dalam menentukan tujuan penelitian sebelum dilakukan proses analisis data.

B. Data Understanding

Pada tahap *Data Understanding*, dilakukan proses pengenalan terhadap dataset yang akan digunakan dalam penelitian. Dataset yang digunakan adalah Groceries Dataset yang diperoleh dari platform Kaggle. Dataset ini berisi riwayat transaksi pembelian pelanggan pada sebuah toko retail yang banyak digunakan sebagai dataset acuan dalam penelitian *Market Basket Analysis* dan *Association Rule Mining* menggunakan algoritma Apriori. Dataset dapat diakses melalui platform Kaggle dengan judul *Groceries Dataset* yang berisi data pembelian produk retail yang sesuai dengan kebutuhan penelitian.

C. Data Preparation

Tahap *Data Preparation* bertujuan menyiapkan dataset agar siap diproses.

Tahapan yang dilakukan meliputi:

- membersihkan data yang kosong (missing value),
- menghapus data duplikat,
- menggabungkan transaksi berdasarkan nomor transaksi,
- melakukan transformasi menggunakan *One Hot Encoding* sehingga data berubah menjadi matriks biner.

D. Modeling

Pada tahap Modeling dilakukan implementasi algoritma Apriori menggunakan library Python. Parameter yang digunakan yaitu :

Minimum Support = 8%

Minimum Confidence = 50%

Minimum Lift = 1

Maximum Rule Length = 3

Algoritma Apriori membentuk *frequent itemset* kemudian menghasilkan *association rule* berdasarkan nilai *support*, *confidence*, dan *lift*. Berikut merupakan rumus-rumus yang akan digunakan.

Rumus *Support* untuk 1 item (SingleItem A):

$$Support(A) = \frac{\Sigma \text{Transaksi mengandung A}}{\text{Total Seluruh Transaksi}} \times 100\% \quad (1)$$

Rumus *Support* untuk kombinasi 2 item (Itemset A dan B):

$$Support(A \cap B) = \frac{\Sigma \text{Transaksi mengandung A dan B}}{\text{Total Seluruh Transaksi}} \times 100\% \quad (2)$$

Rumus *Confidence* untuk aturan $A \rightarrow B$:

$$Confidence(A \rightarrow B) = \frac{\Sigma \text{Transaksi mengandung A dan B}}{\Sigma \text{Transaksi mengandung A}} \times 100\% \quad (3)$$

Atau dapat dinyatakan melalui rasio nilai *support*:

$$Confidence(A \rightarrow B) = \frac{Support(A \cap B)}{Support(A)} \times 100\% \quad (4)$$

Rumus *Lift Ratio* untuk aturan $A \rightarrow B$:

$$Lift(A \rightarrow B) = \frac{Confidence(A \rightarrow B)}{Support(B)} \quad (5)$$

E. Evaluation

Tahap *Evaluation* bertujuan mengevaluasi kualitas aturan asosiasi yang dihasilkan.

Evaluasi dilakukan menggunakan tiga metrik utama yaitu *Support*, *Confidence*, dan *Lift Ratio*. Suatu aturan dianggap layak apabila memenuhi seluruh parameter minimum yang telah ditentukan dan memiliki nilai *Lift* > 1 sehingga menunjukkan adanya hubungan positif antarproduk.

F. Deployment

Tahap Deployment merupakan implementasi hasil model ke dalam aplikasi berbasis Streamlit.

Fitur aplikasi meliputi:

- a) *Upload Dataset*
- b) *Dashboard*
- c) *Frequent Itemset*
- d) *Association Rule*
- e) *Smart Recommendation*
- f) *Insight Bisnis*
- g) *Export PDF/Excel*

Tahap ini menunjukkan bahwa hasil analisis dapat digunakan langsung oleh pengguna sebagai pendukung pengambilan keputusan.

G. Arsitektur Model Sistem

Model aplikasi analisis asosiasi retail otomatis ini dibangun menggunakan arsitektur perangkat lunak modular berbasis *Three-Tier Architecture* (Arsitektur Tiga Lapis) yang memisahkan antara presentasi visual, logika komputasi, dan penyimpanan data. Struktur arsitektur ini dirancang untuk memastikan kinerja aplikasi tetap optimal ketika menangani dataset berskala besar. Komponen arsitektur model tersebut terdiri atas:

a) *Presentation Layer* (Lapisan Antarmuka):

Merupakan lapisan terluar aplikasi yang dikembangkan menggunakan *framework Streamlit Front-End*. Lapisan ini menyediakan antarmuka grafis (GUI) yang interaktif bagi pengguna, seperti widget pengunggah berkas (*st.file_uploader*), penggeser dinamis (*slider*) untuk penentuan parameter support dan *confidence*, representasi grafis *Network Graph* menggunakan pustaka *matplotlib* atau *pyvis*, serta tabel visualisasi data yang interaktif.

b) *Business Logic Layer* (Lapisan Logika Bisnis):

Merupakan inti pemrosesan data (*core processing engine*) yang mengeksekusi algoritma data mining. Lapisan ini ditulis dalam bahasa pemrograman Python dan memanfaatkan pustaka *mlxtend.frequent_patterns* untuk pemrosesan algoritma Apriori (*apriori*) dan pembentukan aturan asosiasi (*association_rules*). Alur logika lapisan ini

bertugas menerima matriks biner hasil pra-pemrosesan, menyaring frequent itemset, dan menghitung nilai metrik asosiasi secara *real-time* berdasarkan masukan parameter dari *Presentation Layer*.

c) *Data Layer* (Lapisan Data):

Merupakan lapisan penyimpanan dan pengelolaan data. Berkas transaksi retail format *.csv* (*Groceries Dataset*) dibaca dan diolah ke dalam struktur data *Pandas Dataframe*. Untuk menjaga efisiensi kinerja dan menghindari proses komputasi berulang, lapisan data ini memanfaatkan fitur *Session State* dari *Streamlit* (*st.session_state*) guna menyimpan dataset bersih dan matriks transaksi biner secara aman di memori server selama sesi pengguna aktif.

III. HASIL DAN PEMBAHASAN

A. *Business Understanding*

Tahap *Business Understanding* pada penelitian ini bertujuan untuk mengetahui pola pembelian pelanggan berdasarkan data transaksi penjualan retail sehingga dapat menghasilkan informasi yang mendukung pengambilan keputusan bisnis. Permasalahan utama yang dihadapi adalah banyaknya data transaksi yang tersimpan, namun belum dimanfaatkan secara optimal untuk memperoleh informasi mengenai keterkaitan antarproduk yang sering dibeli secara bersamaan.

Berdasarkan hasil implementasi algoritma Apriori, tujuan penelitian tersebut berhasil dicapai dengan ditemukannya pola asosiasi antarproduk yang memiliki hubungan signifikan. Sistem mampu mengolah dataset transaksi retail menjadi informasi yang lebih bernilai melalui pembentukan 65 frequent itemset dan 11 *association rule* yang memenuhi parameter *minimum support* sebesar 8%, *minimum confidence* sebesar 50%, dan *minimum lift* sebesar 1. Hasil tersebut menunjukkan bahwa data transaksi yang sebelumnya hanya berfungsi sebagai arsip dapat diubah menjadi informasi yang mendukung proses pengambilan keputusan bisnis.

Informasi yang diperoleh dari aturan asosiasi dapat dimanfaatkan oleh pihak retail dalam menyusun strategi pemasaran, seperti *product bundling*, *cross-selling*, penataan tata letak produk pada rak, serta pengelolaan persediaan barang. Salah satu aturan asosiasi terbaik yang diperoleh adalah kombinasi *other vegetables* dan *rolls/buns* yang mengarah pada pembelian *whole milk*, dengan nilai *support* sebesar 8,21%, *confidence* sebesar 55,94%, dan *lift* sebesar 1,221. Hasil tersebut menunjukkan adanya hubungan positif antarproduk sehingga dapat dijadikan dasar dalam penyusunan strategi promosi yang lebih efektif.

Secara keseluruhan, hasil analisis menunjukkan bahwa penerapan algoritma Apriori mampu memenuhi tujuan bisnis penelitian, yaitu mengubah data transaksi retail menjadi informasi yang dapat digunakan sebagai dasar pengambilan keputusan berbasis data (*data-driven decision making*) untuk meningkatkan efektivitas strategi pemasaran dan operasional pada bisnis retail.

B. Data Understanding

Dataset memiliki tiga atribut utama, yaitu:

Member_number : Nomor identitas pelanggan.

Date/Tanggal : terjadinya transaksi pembelian.

ItemDescription : Nama produk yang dibeli pada transaksi tersebut.

Contoh data transaksi ditunjukkan pada Tabel 1.

Tabel 1. Dataset

Member_Number	Date	Item Description
1808	21/07/2015	tropical fruit
2552	05/01/2015	whole milk
2300	19/09/2015	pip fruit
1187	12/12/2015	other vegetables
3037	01/02/2015	whole milk
4941	14/02/2015	rolls/buns
4501	08/05/2015	other vegetables
3803	23/12/2015	pot plants
2762	20/03/2015	whole milk
4119	12/02/2015	tropical fruit

Dataset awal terdiri atas 38.765 baris data transaksi yang merepresentasikan aktivitas pembelian pelanggan selama periode tertentu. Setiap baris menunjukkan satu item yang dibeli dalam suatu transaksi sehingga satu pelanggan dapat memiliki beberapa baris data pada tanggal transaksi yang sama. Karakteristik tersebut menjadikan dataset sesuai untuk diterapkan pada teknik *Market Basket Analysis*, karena memungkinkan identifikasi pola pembelian produk yang sering muncul secara bersamaan dalam satu transaksi.

Sebelum dilakukan proses pemodelan menggunakan algoritma Apriori, dataset akan melalui tahap Data Preparation untuk memastikan kualitas data dan menyesuaikan format agar dapat diproses oleh algoritma.

C. Data Preparation

Tahapan pemodelan data mining dilakukan dengan memanfaatkan *Groceries Dataset* yang merepresentasikan rekam jejak aktivitas penjualan pada ritel modern. Karakteristik data historis sebelum dan sesudah melewati fase pra-pemrosesan diidentifikasi secara menyeluruh oleh sistem dengan rincian dimensi sebagai berikut:

Tabel 2. Ringkasan Hasil Pra-pemrosesan Data

Keterangan	Jumlah	Deskripsi
Total Baris Awal	38.765 baris	Total rekaman transaksi mentah sebelum proses cleaning.
Total Data Valid	38.006 baris	Baris data setelah penghapusan missing value dan duplikasi produk per pelanggan.
Total Transaksi (Berdasarkan Member)	3.898 transaksi	Total keranjang belanja unik hasil pengelompokan berdasarkan Member number.
Total Produk Unik	167 item	Jumlah jenis barang retail yang tersedia di dalam dataset.

a) Total Baris Data Mentah: 38.765 baris data transaksi terdistribusi dalam basis data tunggal.

b) Total Transaksi Valid (*Cleaned*): 3.898 transaksi pembelian unik setelah melewati perbaikan data duplikat dan *missing values*.

c) Total Kategori Produk Unik: 167 jenis produk retail yang tersebar secara acak di berbagai keranjang belanja konsumen.

D. Modeling

Pada tahap ini, algoritma Apriori diimplementasikan untuk mencari *frequent itemset*, yaitu kombinasi produk yang memenuhi nilai ambang batas tertentu.

Proses ini sangat bergantung pada penentuan parameter metrik utama, yaitu nilai *minimum support* untuk menyeleksi kandidat kombinasi produk yang sering muncul, dan *minimum confidence* untuk menyaring serta memastikan bahwa aturan asosiasi yang terbentuk memiliki tingkat kepastian yang kuat dan valid [3], [7], [8].

Perhitungan matematis untuk mengukur kekuatan aturan asosiasi ($A \rightarrow B$) dilakukan berdasarkan tiga rumus metrik evaluasi berikut [7], [11]:

Nilai Dukungan (*Support*) *Support* adalah metrik yang menunjukkan persentase atau kombinasi kemunculan suatu item atau kombinasi item di dalam keseluruhan basis data transaksi.

Nilai Kepastian (*Confidence*) *Confidence* adalah metrik yang mengukur kuatnya hubungan antar-item dalam aturan asosiasi. Metrik ini menunjukkan kepastian seberapa sering item B dibeli jika pelanggan membeli item A.

Nilai Penguatan (*Lift Ratio*) *Lift Ratio* digunakan sebagai validasi akhir untuk mengukur keandalan dan kekuatan dari *association rule* yang terbentuk. Nilai *lift* menentukan apakah aturan asosiasi tersebut valid dan bermakna secara acak atau tidak.

Pengujian validitas aturan berdasarkan nilai *Lift Ratio* adalah sebagai berikut:

Jika nilai *Lift* > 1 , maka aturan asosiasi dinyatakan valid dan menunjukkan hubungan positif (kemunculan item A sangat memengaruhi pembelian item B).

Jika nilai *Lift* $= 1$, maka item A dan item B bersifat independen (tidak saling memengaruhi). Jika nilai *Lift* < 1 , maka aturan dinyatakan tidak valid atau menunjukkan hubungan negatif.

E. Evaluation

Evaluasi dilakukan untuk mengetahui kualitas aturan asosiasi yang dihasilkan serta memastikan bahwa sistem yang dikembangkan mampu menjalankan seluruh proses analisis sesuai dengan parameter yang telah ditentukan. Pengujian dilakukan menggunakan dataset transaksi retail yang terdiri atas 38.765 baris data, kemudian diproses menjadi 3.898 transaksi valid setelah melalui tahap preprocessing.

Tabel 3. Eksperimen Skenario Parameter Algoritma Apriori

Skenario	Minimum Support	Minimum Confidence	Jumlah Frequent Itemset	Jumlah Rules Terben-tuk	Kesimpulan Analisis
Skenario 1	2% (0.02)	20% (0.20)	876 itemset	1.277 rules	Terlalu banyak aturan, (overfitting), sulit diimplementasikan untuk promosi.
Skenario 2	5% (0.05)	30% (0.30)	245 itemset	154 rules	Jumlah aturan mulai stabil, namun masih banyak asosiasi lemah (lift rendah).
Skenario 3	8% (0.08)	50% (0.50)	65 itemset	11 rules	Sangat Optimal. Menghasilkan aturan asosiasi terkuat dengan lift > 1 yang mudah diinterpretasikan.

Pada tabel 3, Parameter yang paling optimal digunakan pada algoritma Apriori yaitu minimum support sebesar 8%, *minimum confidence* sebesar 50%, minimum *lift* sebesar 1, dan *maximum rule length* sebanyak 3 item.

Berdasarkan hasil pengujian, sistem berhasil menghasilkan 65 *frequent itemset* dan 11 *association rule* yang memenuhi seluruh parameter tersebut. Jumlah aturan yang dihasilkan relatif sedikit dibandingkan apabila menggunakan nilai *minimum support* yang rendah. Hal ini menunjukkan bahwa penggunaan *minimum support* sebesar 8% mampu menyaring kombinasi produk yang kurang signifikan sehingga aturan yang diperoleh lebih relevan dan memiliki peluang lebih besar untuk diterapkan pada lingkungan bisnis retail.

Evaluasi terhadap kualitas aturan asosiasi dilakukan menggunakan tiga metrik utama, yaitu *support*, *confidence*, dan *lift ratio*. Nilai *support* menunjukkan tingkat kemunculan kombinasi produk pada seluruh transaksi, sedangkan *confidence* menggambarkan tingkat keyakinan bahwa suatu produk akan dibeli apabila pelanggan membeli produk tertentu. Sementara itu, *lift ratio* digunakan untuk memastikan bahwa hubungan antarproduk tidak terjadi secara kebetulan, berikut merupakan Tabel 1 menyajikan 10 aturan asosiasi terbaik yang diurutkan berdasarkan nilai *Confidence* dan *Lift* tertinggi.

Tabel 4. Top 10 Aturan Asosiasi Terbaik

No	(Antecedent)	(Consequent)	Support (%)	Confidence (%)	Lift
1	other vegetables	whole milk	8,21	56	1,22
2	shopping bags	whole milk	9,13	54	1,18
3	bottled beer	whole milk	8,54	54	1,17
4	yogurt	whole milk	15,06	53	1,16
5	canned beer	whole milk	8,72	53	1,15
6	bottled water	whole milk	11,24	53	1,15
7	sausage	whole milk	10,7	52	1,13
8	pastry	whole milk	9,11	51	1,12
9	rolls/buns	whole milk	17,86	51	1,11
10	pip fruit	whole milk	8,7	51	1,11

Berdasarkan hasil pada Tabel 4, aturan asosiasi terbaik adalah *other vegetables, rolls/buns* → *whole milk* dengan nilai *Support* sebesar 8,21%, *Confidence* sebesar 55,94%, dan *Lift* sebesar

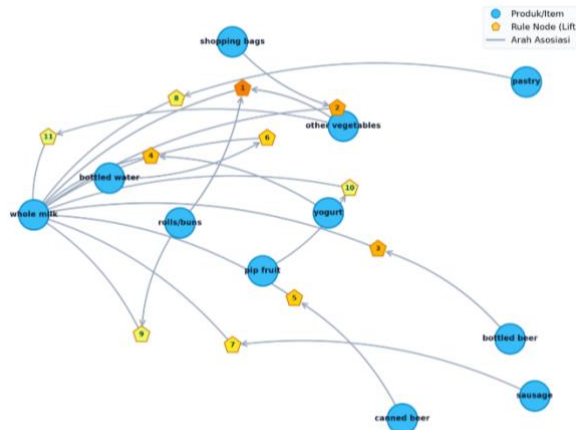
1,221. Nilai *confidence* tersebut menunjukkan bahwa sekitar 55,94% pelanggan yang membeli *other vegetables* dan *rolls/buns* juga membeli *whole milk* dalam transaksi yang sama. Selain itu, nilai *Lift* sebesar 1,221 yang lebih besar dari 1 menunjukkan bahwa hubungan antara ketiga produk tersebut merupakan asosiasi positif dan bukan terjadi secara kebetulan. Dengan kata lain, pelanggan yang membeli *other vegetables* dan *rolls/buns* memiliki kemungkinan sekitar 1,22 kali lebih besar untuk membeli *whole milk* dibandingkan apabila pembelian terjadi secara acak.

Sementara itu, berdasarkan nilai *support* tertinggi, aturan *other vegetables* → *whole milk* memiliki *Support* sebesar 19,14%, yang menunjukkan bahwa kombinasi kedua produk tersebut merupakan pola pembelian yang paling sering muncul pada keseluruhan transaksi pelanggan. Tingginya nilai *support* tersebut mengindikasikan bahwa *whole milk* dan *other vegetables* merupakan dua produk dengan frekuensi pembelian bersamaan yang paling dominan sehingga dapat dijadikan dasar dalam penyusunan strategi *cross-selling*, penataan rak produk, maupun pengelolaan persediaan barang di lingkungan ritel.

Hasil pengujian menunjukkan bahwa seluruh aturan memiliki nilai *lift* lebih besar dari atau sama dengan 1, sehingga dapat dikatakan bahwa seluruh aturan memiliki hubungan asosiasi yang valid. Aturan terbaik diperoleh pada kombinasi *other vegetables* dan *rolls/buns* → *whole milk* dengan nilai *support* sebesar 8,21%, *confidence* sebesar 55,94%, dan *lift* sebesar 1,221. Nilai tersebut menunjukkan bahwa pelanggan yang membeli *other vegetables* dan *rolls/buns* memiliki kecenderungan sekitar 1,22 kali lebih besar untuk membeli *whole milk* dibandingkan pembelian secara acak.

Selain evaluasi terhadap algoritma, dilakukan pula evaluasi terhadap fungsionalitas aplikasi berbasis Streamlit. Seluruh fitur utama, meliputi proses *upload dataset*, *preprocessing data*, pengaturan parameter Apriori, pembentukan

frequent itemset, pembentukan association rule, visualisasi *network graph*, pencarian rekomendasi produk, serta ekspor laporan ke format CSV dan PDF berhasil dijalankan tanpa mengalami kesalahan. Dengan demikian, sistem dinilai telah mampu mendukung proses analisis data transaksi retail secara otomatis dan menghasilkan rekomendasi yang dapat dimanfaatkan sebagai dasar pengambilan keputusan bisnis.

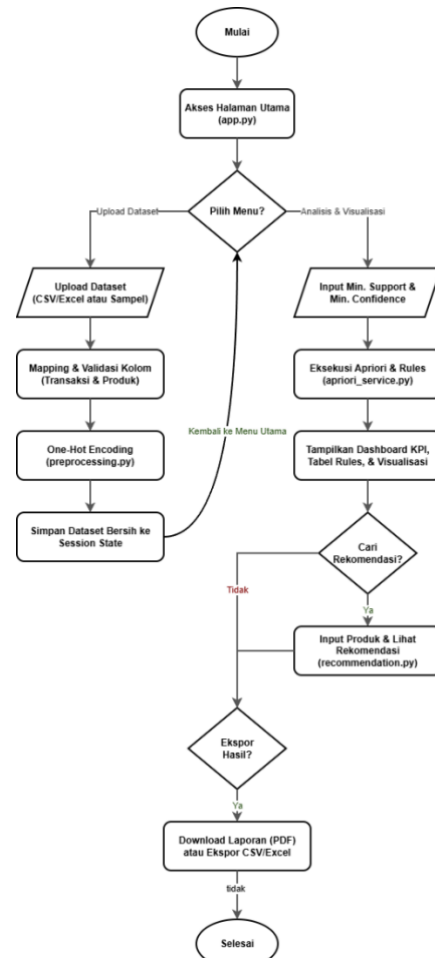


Gambar 2. Visualisasi Network Graph

F. Implementasi

a) Flowchart Sistem

Proses penambangan data (data mining) dari awal pemrosesan hingga menghasilkan luaran rekomendasi bisnis dirancang secara sistematis melalui urutan langkah logis yang terstruktur. Alur data tersebut dapat dijabarkan dalam tahapan-tahapan operasional secara ringkas sebagai berikut:

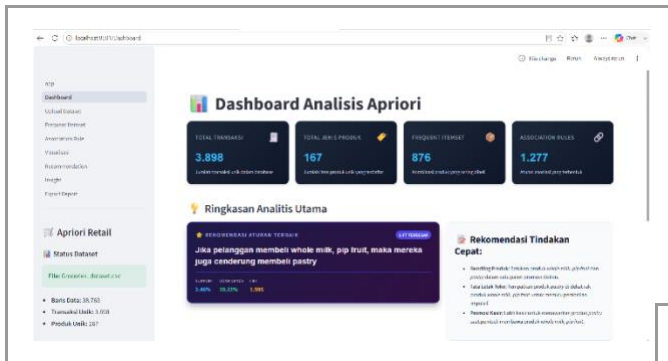


Gambar 3. Visualisasi Flowchart

Implementasi dan Bedah Antarmuka Aplikasi Streamlit

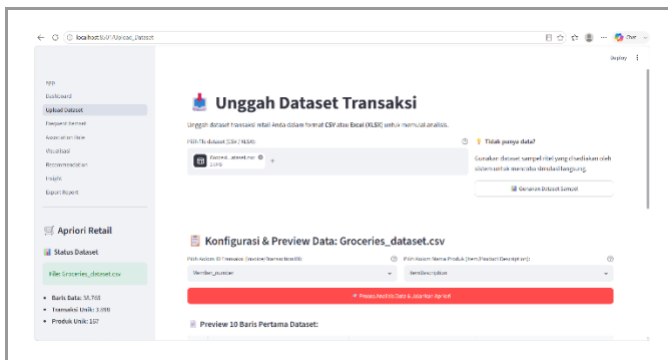
Sistem komputasi otomatisasi algoritma Apriori ini diwujudkan ke dalam platform aplikasi berbasis web interaktif memanfaatkan arsitektur komponen dari *framework* Streamlit. Berdasarkan hasil pengujian sistem, fungsionalitas antarmuka aplikasi terbagi ke dalam 7 modul halaman utama yang dapat diakses melalui menu navigasi *sidebar*:

a) Halaman Dashboard: Menyajikan panel informasi eksekutif ringkas (*executive summary*). Halaman ini secara visual langsung menampilkan tiga metrik utama berupa kartu status dataset, yaitu Total Baris Data (38.765 baris), Total Transaksi Valid (3.898), dan Jumlah Produk Unik (167 item) serta pratinjau grafik frekuensi produk terpopuler.



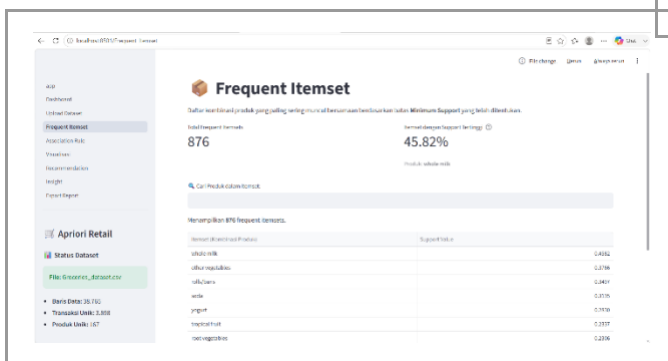
Gambar 4. Tampilan Halaman Dashboard

b) Halaman *Upload Dataset*: Menyediakan komponen *drag-and-drop* untuk memproses file *Groceries_dataset.csv*. Halaman ini dilengkapi pratinjau 10 baris pertama data pasca-transformasi untuk memastikan keaslian struktur data sebelum diproses lebih lanjut oleh mesin algoritma.



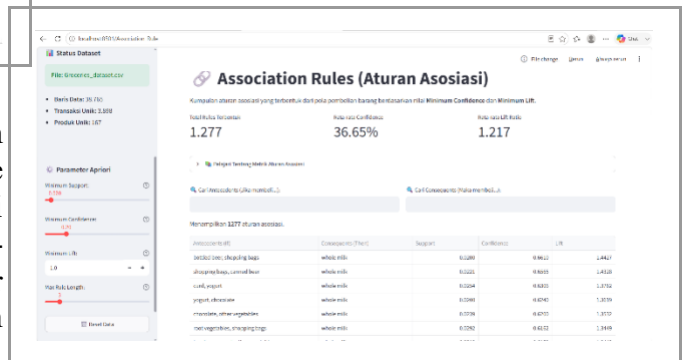
Gambar 5. Tampilan Halaman Dashboard Upload Dataset

c) Halaman *Frequent Itemset*: Menampilkan hasil pemetaan kombinasi barang yang lolos dari seleksi nilai ambang batas *support*. Tabel dilengkapi fitur pengurutan (*sorting*) otomatis untuk mempermudah pencarian kombinasi produk yang memiliki frekuensi kemunculan tertinggi.



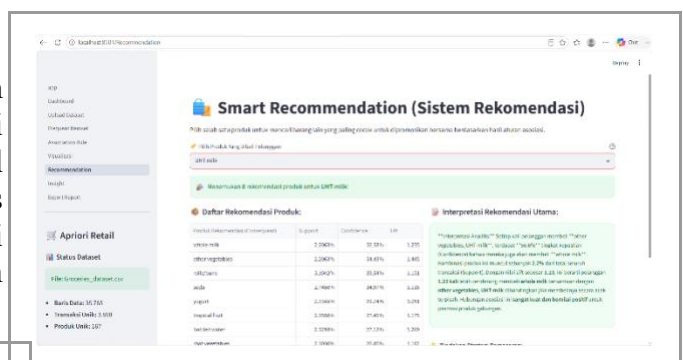
Gambar 6. Tampilan Halaman Frequent Itemset

d) Halaman *Association Rule*: Memuat kumpulan aturan asosiasi dari hasil ekstraksi data transaksi. Pengguna dapat menyaring aturan secara spesifik berdasarkan rentang nilai *confidence* atau memilah produk *antecedent* tertentu melalui kolom pencarian.



Gambar 7. Tampilan Halaman Association Rule

e) Halaman *Smart Recommendation*: Berfungsi sebagai modul sistem rekomendasi pintar. Pengguna cukup memilih satu produk pada menu dropdown (contoh: "yogurt"), dan sistem akan menampilkan visualisasi persentase kecenderungan konsumen untuk membeli produk pasangan (contoh: "whole milk") yang diperkuat dengan instruksi tindakan strategi pemasaran terapan.



Gambar 8. Tampilan Halaman Smart Recommendation

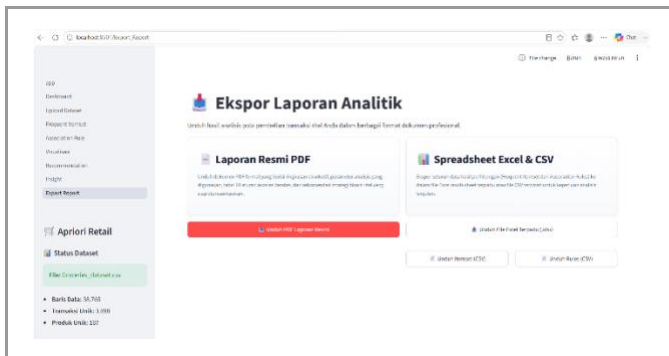
f) Halaman *Insight & Strategi Bisnis*: Menyediakan analisis naratif otomatis (*automated textual insight*) yang digenerasikan oleh sistem berdasarkan aturan asosiasi terkuat yang terdeteksi untuk mendukung pengambilan

keputusan manajemen toko tanpa perlu membaca angka tabel yang rumit.



Gambar 9. Tampilan Halaman Insight

g) Halaman *Export Report*: Menyediakan fasilitas pengunduhan dokumen. Modul ini menyediakan tombol unduh laporan resmi berformat PDF (*Laporan Resmi PDF*) yang rapi serta dokumen kerja Spreadsheet (Excel & CSV) sebagai arsip manajerial perusahaan.



Gambar 10. Tampilan Halaman Export Report

IV. KESIMPULAN

Penelitian ini berhasil mengimplementasikan metode *Market Basket Analysis* menggunakan algoritma Apriori untuk menemukan pola keterkaitan antarproduk berdasarkan data transaksi penjualan ritel. Dataset yang digunakan terdiri atas 38.765 baris data, 3.898 transaksi valid, dan 167 jenis produk. Dengan menggunakan parameter *minimum support* sebesar 8%, *minimum confidence* sebesar 50%, *minimum lift* sebesar 1, serta *maximum rule length* sebanyak 3 item, sistem berhasil menghasilkan 65 *frequent itemset* dan 11 *association rules* yang memenuhi seluruh kriteria yang telah ditentukan. Hasil ini menunjukkan bahwa penggunaan nilai support yang lebih tinggi mampu menyaring aturan asosiasi sehingga

hanya menghasilkan pola pembelian yang benar-benar sering muncul dan memiliki hubungan yang signifikan.

Berdasarkan hasil analisis, aturan asosiasi terbaik diperoleh pada kombinasi *other vegetables* dan *rolls/buns* yang mengarah pada pembelian *whole milk*, dengan nilai *support* sebesar 8,21%, *confidence* sebesar 55,94%, dan *lift* sebesar 1,221. Nilai *confidence* tersebut menunjukkan bahwa lebih dari separuh pelanggan yang membeli *other vegetables* dan *rolls/buns* juga membeli *whole milk*, sedangkan nilai lift yang lebih besar dari satu mengindikasikan adanya hubungan asosiasi positif antarkategori produk. Selain itu, aturan *other vegetables* → *whole milk* memiliki nilai *support* tertinggi sebesar 19,14%, sehingga menjadi pola pembelian yang paling sering muncul pada keseluruhan transaksi.

Selain menghasilkan aturan asosiasi, penelitian ini juga berhasil mengembangkan sebuah aplikasi berbasis web menggunakan Streamlit yang mampu mendukung proses analisis data secara interaktif. Aplikasi menyediakan fitur pengaturan parameter Apriori, visualisasi hubungan antarproduk dalam bentuk network graph, pencarian rekomendasi produk berdasarkan aturan asosiasi, serta fasilitas ekspor hasil analisis ke dalam format CSV dan PDF. Fitur-fitur tersebut memudahkan pengguna dalam melakukan analisis maupun mendokumentasikan hasil penambangan data.

Dari sisi bisnis, informasi yang dihasilkan dapat dimanfaatkan sebagai dasar dalam penyusunan strategi product bundling, cross-selling, penataan ulang tata letak produk, serta pengelolaan persediaan barang. Dengan memanfaatkan pola pembelian pelanggan yang diperoleh dari algoritma Apriori, pihak ritel dapat meningkatkan efektivitas promosi, mengoptimalkan ketersediaan stok, dan mendukung pengambilan keputusan yang lebih tepat berdasarkan data transaksi yang dimiliki.

REFERENSI

- [1] A. Wijaya, A. Faqih, D. Solihudin, C. L. Rohmat, and S. E. Permana, "Penerapan Association Rules Menggunakan Algoritma Apriori Untuk Identifikasi Pola Pembelian," 2023.
- [2] D. Laila Safitry *et al.*, "Analisis Pola Pembelian Konsumen Menggunakan Algoritma Apriori Untuk Menentukan Strategi Pemasaran Produk Di Toko Retail X," 2025.
- [3] I. Utnasar, "Analisis Data Pola Pembelian Konsumen Dengan Algoritma Apriori Pada Transaksi Penjualan Minimarket D Mart," *Jurnal Sistem Informasi Dan Informatika*, vol. 2, no. 1, 2024, [Online]. Available: <http://jurnal.unidha.ac.id/index.php/jteksis>
- [4] I. F. Rahman and D. Riana, "Market Basket Analysis untuk Penjualan Retail: Perbandingan Akurasi Algoritma Apriori dan FP-Growth Berbasis CRISP-DM," *Jurnal Algoritma*, vol. 22, no. 1, pp. 468–479, May 2025, doi: 10.33364/algoritma/v.22-1.2303.
- [5] A. N. Aulia, E. Mufida, L. K. Rahayu, E. Rachmawati, H. Hertyana, and D. Andriansyah, "Analisa Pola Asosiasi Produk Pada Transaksi Penjualan Retail Online Menggunakan Algoritma Apriori," *Jurnal INSAN (Journal of Information Systems Management Innovation)*, vol. 6, no. 1, 2026, [Online]. Available: <http://jurnal.bsi.ac.id/index.php/jinsan>
- [6] N. M. Mustofa, A. M. Alfarisi, and A. Tholib, "Penerapan Algoritma Apriori Untuk Menentukan Pola Pembelian Konsumen," *JSITIK: Jurnal Sistem Informasi dan Teknologi Informasi Komputer*, vol. 4, no. 1, pp. 11–30, Dec. 2025, doi: 10.53624/jsitik.v4i1.710.
- [7] S. Aulia Miranda, F. Fahrullah, and D. Kurniawan, "Implementasi Association Rule Dalam Menganalisis Data Penjualan Sheshop dengan Menggunakan Algoritma Apriori," *METIK JURNAL*, vol. 6, no. 1, pp. 30–36, Jul. 2022, doi: 10.47002/metik.v6i1.342.
- [8] R. Suganda and A. Solichin, "Analisis Transaksi Penjualan Produk Minuman Kopi Menggunakan Metode Association Rule Dengan Algoritma Apriori Analysis of Sales Transactions for Coffee Beverage Products Using The Association Rule Method With The Apriori Algorithm," *KRESNA: Jurnal Riset dan Pengabdian Masyarakat*, vol. 4, no. 1, 2024, [Online]. Available: <https://jurnaldrpm.budiluhur.ac.id/index.php/Kresna/>
- [9] S. Butsianto, Candra Naya, and Anggi Muhammad Rifa'i, "Implementasi Algoritma Apriori dalam Menemukan Pola Asosiasi pada Data Penjualan Produk Retail," *Bulletin of Computer Science Research*, vol. 5, no. 5, pp. 938–947, Aug. 2025, doi: 10.47065/bulletincsr.v5i5.731.
- [10] D. Selvida *et al.*, "Nearest Neighbor untuk Analisis Pola Prediksi Pembelian," *Journal of Informatics, Information System, and Artificial Intelligence*, vol. 4, no. 1, pp. 54–63, 2026.
- [11] Septianingsih and A. Bayu Santoso, "Implementasi Association Rule Mining Menggunakan Algoritma Apriori Untuk Rekomendasi Cross-Selling Produk Ritel," *Jurnal Algoritma*, vol. 23, no. 1, May 2026, doi: 10.33364/algoritma/v.23-1.3381.
- [12] L. M. Zidan, "Jurnal Komisi (Komputer dan Sistem Informasi) Analisis Aturan Asosiasi pada Data Transaksi," *Jurnal Komisi (Komputer dan Sistem Informasi)*, vol. 2, no. 2, pp. 41–50, 2025, [Online]. Available: <https://komisijournal.indiepress.id/index.php/komisi/index>
- [13] H. R. Wijaya, K. Khairil, and D. Lianda, "Implementasi Data Mining Menggunakan Algoritma Apriori dalam Menentukan Persediaan Barang," *Digital Transformation Technology*, vol. 5, no. 1, pp. 427–433, Jul. 2025, doi: 10.47709/digitech.v5i1.6381.
- [14] S. D. Rahmawati, A. B. Oktavia, F. S. A. Putri, and D. L. Fithri, "Penerapan Algoritma Apriori Untuk Menemukan Pola Asosiasi Pada Data Penjualan Retail Fashion," *SIMKOM*, vol. 10, no. 2, pp. 289–295, Jul. 2025, doi: 10.51717/simkom.v10i2.910.
- [15] Rizka, H. Pratama, P. Nabawy, B. Cahyadi, and Mhd. Furqan, "Analisis Pola Asosiasi Interaksi Pengguna pada Sistem Informasi Akademik Berbasis Web Menggunakan Algoritma Apriori," *Data Sciences Indonesia (DSI)*, vol. 5, no. 1, pp. 10–17, Jul. 2025, doi: 10.47709/dsi.v5i1.5943.